SLIDING WINDOW AUTOCORRELATION FOR THE SYNCHRONOUS UNSUPERVISED SEGMENTATION OF THE PHONOCARDIOGRAM

Sankua Chao and Adrian D. C. Chan
Department of Systems and Computer Engineering, Carleton University, Ottawa, Ontario

ABSTRACT

A wavelet pre-processor is examined as an extension to the Sliding Window Autocorrelation (SWA) phonocardiogram (PCG) segmentation algorithm. The method by which the SWA performs unsupervised, synchronous segmentation of heartbeats in a PCG is described. The appropriateness of using Morlet wavelets for pre-processing PCGs is also described. A possible implementation of a wavelet pre-processor is presented with preliminary experimental results.

INTRODUCTION

The phonocardiogram (PCG) is a bio-signal that represents sounds produced by the heart. During a typical heartbeat, the first heart sound (S1) occurs at the beginning of systole, and the second heart sound (S2) occurs at the beginning of diastole. S1 is caused by the closing of the atrial-ventricular valves (i.e. mitral and tricuspid), whereas S2 is caused by the closing of the semilunar valves (i.e. aortic and pulmonary) [1]. The PCG can be used to detect and diagnose murmurs, which are usually associated with heart valve malfunctions (e.g. stenosis and regurgitation) [1].

The long-term goal of this work is to develop an automated PCG analysis system. Automatic murmur detection and diagnoses are well researched. Prior to murmur detection, it is necessary to segment the PCG to obtain timing information. PCG segmentation also allows alignment of multiple segments of the cyclostationary PCG signal, improving the signal to noise ratio of the PCG during analysis. The majority of PCG research has focused on event detection assuming accurate PCG segmentation, or has used the electrocardiogram (ECG) for PCG segmentation. This work investigates PCG segmentation that does not require the ECG, hence simplifying the patient interface in a PCG analysis system.

In this work, a wavelet pre-processor is examined as an extension to the Sliding Window Autocorrelation PCG segmentation algorithm. This pre-processing seeks to isolate relevant heart sounds (i.e. S1 and S2), suppressing noises not originating from the heart. Wavelets have been shown to be effective tools for PCG segmentation, in conjunction with techniques such as correlator banks [2][3], Shannon energy [4], and complexity adaptive filters [5]. In conjunction with averaging techniques, wavelets have also been used for denoising PCG signals [6]. It is anticipated that wavelets will prove useful for PCG segmentation.

A wavelet transform can provide an indication of the time intervals in a PCG signal at which different frequencies occur. Since S1 and S2 have been shown to have specific frequency ranges (10-180 Hz and 50-250 Hz, respectively [7]), analysis of the wavelet transform can indicate when S1 and S2 likely occur in a PCG signal.

SLIDING WINDOW AUTOCORRELATION

The Sliding Window Autocorrelation (SWA) is a proposed PCG segmentation algorithm, which is designed to perform unsupervised, synchronous segmentation of complex heart sounds [8].

A Heart Rate Estimator (HRE) is used which, in conjunction with *a priori* patient information, provides an estimate of the heartbeat period T_{beat}; the estimated heartbeat period is used to set particular parameters of the SWA. The estimated heartbeat period corresponds to the first large peak of the PCG signal energy's autocorrelation. The time offset range corresponding to a reasonable heartbeat period is emphasized, by multiplying the autocorrelation by a tapered window.

Central to the performance of the SWA is the calculation of the *similarity correlation* $G(\tau)$ using the normalized PCG signal magnitude $w(t)$. Similar to a cross-correlation, the time offset at which a peak occurs indicates the probable time difference between two highly similar windows. The similarity correlation, however, attempts to compensate for beat-to-beat magnitude variations, by scaling the cross-correlation by the magnitudes of each window.

Each element of $G(\tau)$ is calculated using Equations (1)-(4). The two windows to be correlated, $w_1(t)$ and $w_2(t)$, are of length T_1 and T_2, respectively (by convention $T_1 \le T_2$). The conventional cross-

correlation $R(\tau)$ involves shifting $w_1(t)$ by τ, where $0 \le \tau < T_2$, as shown in (1). The similarity correlation $G(\tau)$ is scaled by magnitudes derived from (2) and (3).

$$R(\tau) = \sum_{t=0}^{T_2} w_1(t-\tau) w_2(t) \quad (1)$$

$$B = \sum_{t=0}^{T_1} [w_1(t)]^2 \quad (2)$$

$$C(t) = A(t) \bullet N(t) \ , \text{ where:} \quad (3)$$

$A(t) = [w_2(t)]^2$, for $0 \le t < T_2$;

$N(t) = 1$, for $0 \le t < T_1$;

and $\bullet$ denotes cross-correlation.

$$G(\tau) = \frac{R(\tau)}{\left|\sqrt{B}\right|\left|\sqrt{C(\tau)}\right|} \quad , \quad 0 \le \tau < T_2 \quad (4)$$

The SWA performs two sets of similarity correlations. For the first set of similarity correlations, a similarity correlation is performed at each 40 ms time step in the PCG signal. Each similarity correlation is generated by sliding a short window (representing 75% of T_{beat}) across a longer window. The time offsets of the peaks, when taken in sequence, are referred to as the *offset sequence*. When viewed graphically, as illustrated in Figure 1, the longest portion of the offset sequence that maintains a relatively constant magnitude provides a rough indication of where the *heartbeat template* starts. The heartbeat template is central to the second set of similarity correlations.

For the second set of similarity correlations, each similarity correlation is generated by sliding the heartbeat template across a longer window. Similarity correlations are performed towards the right and left of the heartbeat template. Towards the right of the heartbeat template, the first longer window is positioned between 60%–180% of T_{beat} away from the heartbeat template. Towards the left of the heartbeat template, the first longer window is positioned between 0%–150% of T_{beat} away from the heartbeat template. The time offset of a peak corresponds to the time difference to the neighbouring heartbeat. These time differences, along with knowledge of the actual position of the heartbeat template within the PCG signal, are used to calculate the times at which heartbeats occur, which correspond to the boundaries between heartbeats. In this manner, the PCG signal is segmented. Illustrated in Figure 2 is the second set of similarity correlations and predicted heartbeat boundaries.

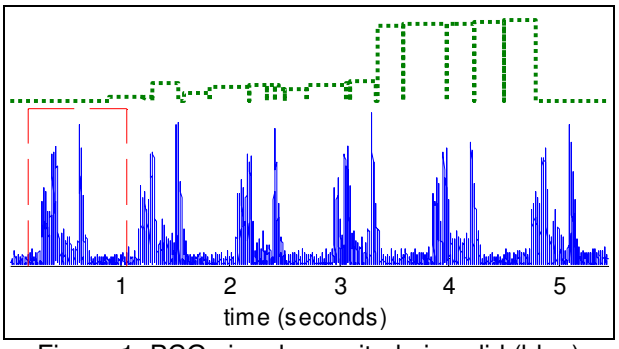

Figure 1: PCG signal magnitude is solid (blue) waveform. Offset sequence is dotted (green) waveform. Longest portion of offset sequence that maintains constant magnitude is located between 0 to 1 second. Heartbeat template indicated by dashed (red) box.

The SWA performs synchronous segmentation, in the sense that heartbeat boundaries are located a consistent time before S1 in a given PCG signal. For the purpose of this work, it is not essential to represent the beginning of a heartbeat as the start of systole (i.e. S1), as the primary concern is the alignment of multiple PCG segments.

Original results generated by the SWA in [8] were evaluated using several conventions. A correct boundary should be located at a time before systole (i.e. S1) that is consistent with other boundaries in the PCG signal. An incorrect boundary could be located at a distance before systole that is inconsistent with other boundaries in the PCG signal, or could be located during systole (i.e. between S1 and S2).

On a dataset containing PCG signals classified as *simple* (minimal diastolic noise, allowing for distinct heartbeat edges), *moderate* (murmur noise, causing

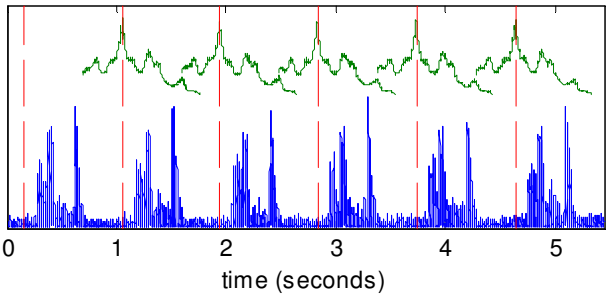

Figure 2: PCG signal magnitude is solid (blue) waveform in lower half. Second set of similarity correlations represented by solid (green) waveforms in upper half. Predicted heartbeat boundaries indicated by dashed (red) lines.

indistinct heartbeat edges), and *complex* (non-cardiologic noise, such that relevant heart sounds are difficult to hear or visualize, as well as highly irregular heartbeat sequences), the SWA achieved an overall accuracy of 83.2%. The accuracy of the SWA for each subset of PCG signals can be seen in Table 1. The accuracy of the SWA is limited by the HRE, which achieved an overall accuracy of 90.8% [8].

WAVELET PRE-PROCESSOR

A wavelet pre-processor is expected to better isolate the information to be correlated by the SWA. When a heart sound is masked by other noises (such as murmurs, talking, crying, etc.), the location of the heart sound is uncertain. Consequently, it is more difficult for the SWA to identify heartbeat boundaries when masked heart sounds are correlated. Wavelets could be used to remove confounding noises from relevant heart sounds, and hence improve the signal to noise ratio of a PCG signal prior to segmentation.

In a PCG signal, heart sounds are represented as short time pulses, and a time-frequency analysis would be beneficial. The analysis should reveal knowledge of what frequencies occur in a PCG signal, as well as when the frequencies occur in a PCG signal. Conventional Fourier methods involve only frequency analysis, and are not well-suited for this work as they provide no time localization. Wavelets can discern the frequency characteristics of heart sounds from noise, as well as localize the frequency characteristics in time; and, hence, are anticipated to be an appropriate tool for analyzing PCG signals.

In general, a wavelet transform involves multiplying and integrating a wavelet with each time interval of a PCG signal, and repeating for scaled versions of the wavelet. A lower scale value generates a more compact or compressed wavelet (i.e. narrowing width of central oscillation), which can correlate well with small periods associated with high frequencies. Hence, using a low scale value allows a wavelet to detect high frequencies in a PCG signal.

The result of performing a wavelet transform on a PCG signal is a matrix of coefficient values, where one coefficient ($C_{f,t}$) exists for each scale (or corresponding frequency f) at each time interval t of a PCG signal. The magnitude of a coefficient indicates the presence of sound peaks of frequency f at time interval t. Hence, a wavelet transform can be used to determine when sound peaks of particular frequencies (i.e. corresponding to relevant heart sounds) occur in a PCG signal.

Table 1: Prediction accuracy of SWA segmentation.

Dataset	Number of heartbeats	Accuracy
Simple	1948	88.9 %
Moderate	589	83.7 %
Complex	172	16.9 %
Overall	2709	83.2 %

The type of wavelet used for performing a wavelet transform should meet several criteria. The wavelet should be compact enough to correlate with the highest possible frequency in a PCG signal. The wavelet should be scaleable, in order to correlate with the various frequencies in a PCG signal. The shape of the wavelet should be similar to the relevant heart sounds. Based on the rationale presented in previous work [2], the Morlet wavelet will be initially investigated for the wavelet pre-processor in this work. The Morlet wavelet contains multiple oscillations, resembling the oscillatory nature of the relevant heart sounds [2]. The Morlet is also scaleable, so it can correlate with both S1 and S2.

One possible implementation of the wavelet pre-processor uses a threshold technique. Using the Morlet wavelet, a wavelet transform is performed on the original PCG signal $s_{orig}(t)$. Coefficients with an absolute value above a particular threshold indicate the time intervals containing sound peaks. Portions of $s_{orig}(t)$ outside of these time intervals are nulled, resulting in a modified PCG signal $s_{post}(t)$ in which only sound peaks are present. To perform PCG segmentation in conjunction with wavelet pre-processing, the SWA would be applied to $s_{post}(t)$.

PRELIMINARY RESULTS

Wavelet pre-processing was applied to PCG signals from the dataset of [8], using the threshold technique described in the previous section. Preliminary results for a PCG signal from the simple and moderate subset are presented here.

The simple PCG signal [9] is representative of aortic valve stenosis (insufficient closure of aortic valve, during diastole) with ejection click (abrupt opening of semilunar valves, early in systole [10]). The original PCG signal, coefficients of the wavelet transform, and two modified PCG signals (subjected to low and high threshold) are shown in Figure 3.

The moderate PCG signal [11] is representative of tetralogy of fallot (four congenital heart defects [10], causing murmurs during systole). The original PCG signal, coefficients of the wavelet transform, and two modified PCG signals (subjected to low and high threshold) are shown in Figure 4.

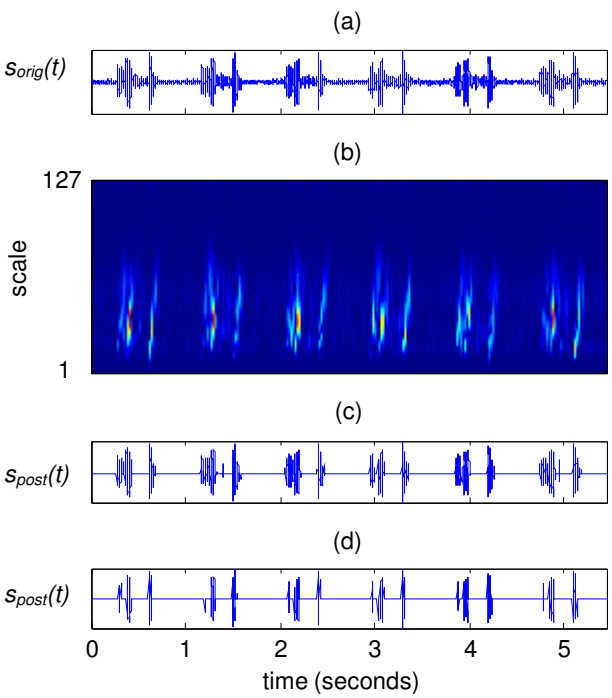

Figure 3: (a) Simple PCG signal. (b) Wavelet transform coefficients. Modified signal subjected to (c) low threshold, and (d) high threshold.

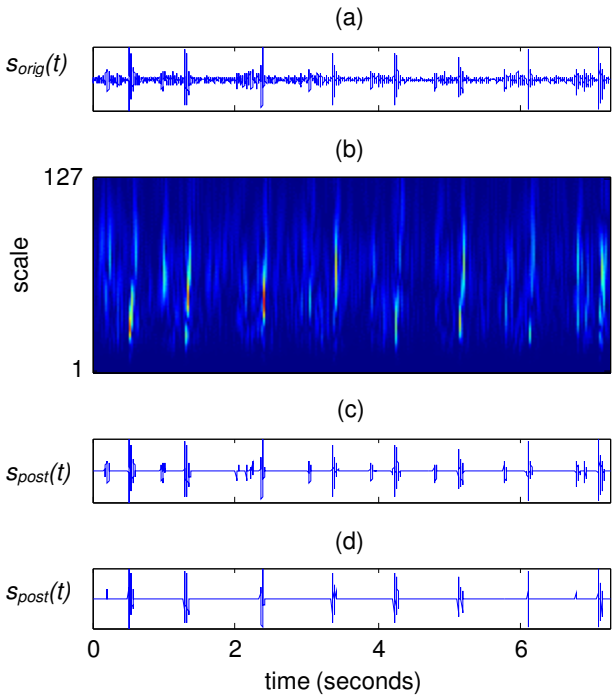

Figure 4: (a) Moderate PCG signal. (b) Wavelet transform coefficients. Modified signal subjected to (c) low threshold, and (d) high threshold.

Wavelet denoising has demonstrated the ability to remove a significant amount of the noise. A high threshold can cause more of $s_{orig}(t)$ to be nulled, such that S1 or S2 may not be present in $s_{post}(t)$ (Figure 4d). Such distortion may be acceptable, since this work does not attempt to characterize or classify individual heart sounds, but rather attempts to simply extract timing information for PCG segmentation.

Future implementations of the wavelet pre-processor could use a different threshold technique, where frequency bands beyond the frequency ranges of relevant hearts sounds will be excluded from PCG reconstruction.

CONCLUSION

The use of a wavelet pre-processor is expected to improve the signal to noise ratio of PCG signals, and subsequently improve the overall performance of the SWA segmentation algorithm. Preliminary results from one possible implementation of the wavelet pre-processor, using a threshold technique and a Morlet wavelet, show the successful isolation of heart sound peaks in a simple and a moderate PCG signal. Further research will verify and quantify the improved performance of the SWA segmentation algorithm when extended by a wavelet pre-processor.

REFERENCES

[1] J.G. Webster, *Medical Instrumentation: Application and Design*, 3rd Edition, John Wiley & Sons, Hoboken, NJ, 1998.

[2] S. Rajan, "A wavelet-based correlator approach for segmentation and classification of the first two heart sounds in a PCG signal," Ph.D. dissertation, The University of New Brunswick, NB, 2004.

[3] S. Rajan, E. Budd, M. Stevenson, and R. Doraiswami, "Unsupervised and uncued segmentation of the fundamental heart sounds in phonocardiograms using a time-scale representation," *Proc. 28th Ann. Intl. Conf. IEEE EMBS*, pp. 3732-3735, 2006.

[4] P. Wang, Y. Kim, L.H. Ling, and C.B. Soh, "First heart sound detection for phonocardiogram segmentation," *Proc. 27th Ann. Intl. Conf. IEEE EMBS*, pp. 5519-5522, 2005.

[5] D. Kumar, P. Carvalho, M. Antunes, J. Henriques, M. Maldonado, and R. Schmidt, "Wavelet transform and simplicity based heart murmur segmentation," *Computers in Cardiology*, vol. 33, pp. 173-176, 2006.

[6] S.R. Messer, J. Agzarian, and D. Abbott, "Optimal wavelet denoising for phonocardiograms," *Microelectronics Journal*, vol. 32, pp. 931-941, 2001.

[7] S.M. Debbal and F. Bereksi-Reguig, "Spectral analysis of the PCG signals," *The Internet Journal of Medical Technology*, vol. 4, no. 1, 2007.

[8] P. Beirne, "Unsupervised segmentation of heart sounds," M.A.Sc. dissertation, Carleton University, Ottawa, ON, 2006.

[9] C.A. Altman, M.R. Nihill, and J.T. Bricker, *Pediatric Cardiac Auscultation*, Lippincott Williams & Wilkins, 2000.

[10] H.K. Walker, W.D. Hall, and J.W. Hurst, *Clinical Methods: The History, Physical, and Laboratory Examinations*, Butterworth, 1990.

[11] D.L. Roy and B. Hoyt, *EarsOn!*, Cor Sonics Inc., Halifax, NS, 1997.